

RUNNING HEAD: Is PAR a Good Investment?

Is PAR a Good Investment?

Understanding the Costs and Benefits of Teacher Peer Assistance and Review Programs

John P. Papay
Research Assistant

Susan Moore Johnson
Jerome T. Murphy Professor in Education

Harvard Graduate School of Education
Project on the Next Generation of Teachers

January 2011

Abstract

Peer Assistance and Review (PAR) is a local labor–management initiative designed to improve teacher quality. In PAR, expert “consulting teachers” mentor, support, and evaluate novice and underperforming veteran teachers. Evaluations under PAR can lead to dismissals. We examine the costs and benefits of PAR, both financial and organizational. Although PAR is an expensive reform, costing \$4,000 to \$10,000 per teacher served, it affords the district a range of financial savings and organizational benefits that offset program costs. We argue that limiting the scope of an educational cost-benefit study to only quantifiable elements artificially constrains understanding what a reform actually requires and offers.

Keywords: Evaluation, Teacher assessment, Teacher research, Educational reform

Is PAR a Good Investment?
Understanding the Costs and Benefits of Teacher Peer Assistance and Review Programs

With evidence mounting that the quality of teachers' work is crucial to students' learning, policymakers and school district officials have intensified the search for ways to attract, support and retain effective teachers. Many observers argue that developing high-quality professional evaluation systems is a necessary step towards improving the U.S. teaching force (Weingarten, 2010; Donaldson, 2009). However, administrators' evaluations of teachers are infrequent and superficial in most districts, rarely leading to deliberate decisions about tenure or dismissals for poor performance. Often, teachers report that they go several years without principals observing them for their "annual" evaluations, and when such assessments do occur they are typically "drive-by" evaluations based on five-to-ten minute class visits (Toch & Rothman, 2008). Thus, very few teachers are judged to be ineffective and even fewer are ever dismissed (Weisberg et al., 2009; Dawson & Billingsley, 2000; Tucker, 1997; Ward, 1995).

One reform for improving teacher quality – Peer Assistance and Review (PAR) – offers an alternative approach to these cursory evaluation systems. PAR focuses not only on supporting and assessing individual teachers, but also on expanding the capacity of the school and district to improve teaching and learning. Recently, many educational observers and policymakers, including President Barack Obama, Secretary of Education Arne Duncan, and American Federation of Teachers President Randi Weingarten, have pointed to PAR as an approach with great potential for improving professional evaluation and teacher quality (Obama, 2009; Duncan, 2009; Dillon, 2008; Toch & Rothman, 2008; Goldstein, 2007; Koppich, 2005; Goldstein, 2004).

PAR, which began in Toledo, OH nearly thirty years ago, has been adopted by several dozen school districts across the country. Districts with PAR appoint expert teachers—often

called “consulting teachers” (CTs)—to mentor, support, and evaluate teachers in the program. PAR typically includes a Novice program, which provides induction for all new teachers in the district and makes renewal recommendations. Most districts also have an Intervention program that serves underperforming veteran teachers, identified by principals as not meeting the district’s instructional standards through the regular evaluation process. If these teachers do not improve to meet standards, PAR can lead to the dismissal of tenured teachers. All other teachers in the district are evaluated by principals in the traditional manner, although some districts, like Cincinnati, have developed systems that incorporate peer evaluations for all teachers.

Consulting teachers are typically selected through a competitive process. They usually leave the classroom for several years, but most districts require them to return to full-time teaching after a three-year term. In their role, CTs are assigned a caseload of ten to twenty teachers in PAR. In most districts, CTs observe and meet with each teacher at least once a week. They document thoroughly what they see during the lesson and assess these observed practices using the district’s instructional standards. They also typically have a debriefing meeting with the teacher to discuss progress, to identify areas for improvement, and to offer support and advice.

At least twice a year, CTs also write a more complete summative evaluation of the teacher’s performance, which they present to the joint labor-management PAR Panel that governs the program. The Panel’s primary responsibility is to review these cases and to decide whether the district should continue to employ those teachers based on the evidence collected by the CTs. Panel members listen to these presentations, question the CTs, review the evidence, and eventually decide whether to recommend that the district retain or dismiss the teachers. In almost

all cases, if the CT judges that a teacher has not met the district's instructional standards, that teacher faces dismissal.¹

PAR thus has several advantages over traditional teacher evaluation systems. CTs conduct regular and thorough evaluations, enabling them to offer consistent support and to assess the teacher's performance and document her progress. With such support, PAR provides strong induction for novice teachers and intensive, directed professional development for underperforming veterans. However, if teachers in either group cannot meet the district's standards with this support, they can be dismissed with the oversight and agreement of both the teachers union and district management.

This intensive commitment of human resources to evaluating teachers in PAR makes it an expensive program. However, to date, no research has systematically examined the program's costs or its benefits. In this study, we build on traditional cost analysis approaches to assess the costs of PAR, and we document the program's benefits. We first describe our analytic approach as well as the design of our study of PAR programs in seven districts across the country. We then examine the financial costs and short-term administrative cost-savings that result from implementing PAR. However, any analysis of the costs and benefits of PAR – a program designed to improve instructional quality – should also look beyond the short-term monetary trade-offs to examine broader outcomes like student achievement. Unfortunately, the value of PAR's effects on student achievement is quite difficult to quantify, both because there are no estimates of PAR's impact on student achievement and because monetizing any such benefits requires relying on large assumptions. In this paper, we present one approach to account for possible effects in a cost-benefit framework. Finally, one of the most striking findings from our research is how important other factors are to the stakeholders in these districts. Thus we

¹ See Goldstein, 2007, Author, 2009, and Author, 2010, for more detailed descriptions of PAR programs.

examine PAR's broader, organizational benefits, which can improve the capacity of schools to serve all students, even though their effects are not quantifiable in the short-term.

The current economic context has pushed many districts to assess the financial costs and benefits of reforms narrowly and in a short-term time frame. However, we argue that, even—and perhaps especially—in a time of particularly scarce resources for many districts, taking a broader perspective on the costs and benefits of PAR, or any program, is both warranted and wise. In time, researchers may be able to develop rigorous causal estimates of PAR's long-term effects on student learning; however, for now, policymakers must decide whether or not to invest in the program. Our evidence suggests that PAR is likely to be a worthwhile investment for many districts because it reduces administrative costs and improves the quality of teaching.

Section II. Data and Methods

Approach to Cost-Benefit Analysis

Despite the growing interest in educational program evaluation, analyses that consider both program effects and their costs remain rare in educational research (Hummel-Rossi & Ashdown, 2002; Rice, 2002). In a recent review of cost studies in education, Ross, Barkaoui, & Scott (2007) focused on quantitative work that compares expenditures to estimates of program benefits, typically through random assignment experiments or quasi-experimental analyses. They conclude that “educational evaluators have shown little interest in addressing questions about whether the programs they evaluate are worth the money these programs cost” (p. 477).

Cost analyses typically follow several steps (Boardman et al., 2006; Levin & McEwan, 2001; Levin, 1983). First, the program's financial costs are enumerated. Costs include not only program expenditures but the full value of all resources used in the program. Second, program

benefits are assessed, typically through a causal evaluation of program effectiveness. Third, these program benefits are converted to a common metric, such as dollars in a cost-benefit analysis.² Monetizing program benefits requires analysts to make assumptions about the values of specific program outcomes. Finally, program benefits are compared to program costs. To the extent that program benefits outweigh costs, a program is deemed to be cost-effective.

There are at least three important challenges to using traditional cost-benefit analyses for a complex educational intervention like Peer Assistance and Review. First, it is necessary to include all of the possible outputs of such a program, but these benefits are often spread widely and hard to assess. As Rice writes, “although it is often difficult to identify and assign a value to many of the inputs and outputs of educational interventions, their inclusion in the total cost and effectiveness estimates is essential” (1997, p. 311). To include all of these factors in the analysis, researchers must develop estimates of the program’s impact on all possible outcomes. For the analysis to be credible, these estimates must derive from a well-designed causal impact evaluation (McEwan, 2002).

However, developing unbiased causal estimates proves challenging even in the case of small, uniformly implemented interventions. For example, evaluating whether a program improves student achievement is difficult because districts choose whether or not to adopt the program. Thus, any subsequent differences in student achievement could be the result of the program or they could simply reflect other pre-existing differences between districts. This challenge has led to the prominence of randomized control trials in educational research, but such studies are expensive and challenging to implement in their own right. Evaluating reforms

² “Cost-effectiveness” analysis takes a more flexible approach, with outcome measures evaluated in their natural units. However, in practice, cost-effectiveness approaches attempt to evaluate programs using a single, common outcome (e.g., student test scores) and compare the benefits to program costs.

like PAR that are meant to change an entire district's practices and necessarily vary widely in implementation from place to place is much more difficult.

In the case of PAR, rigorous impact evaluation is complicated by a number of factors. First, relatively few sites have adopted the program, so the sample size for an evaluation is limited. Second, the decision to implement PAR is clearly a choice that districts make. Thus, researchers cannot simply compare student outcomes in districts that have PAR and similar districts that do not. Third, PAR is frequently implemented at the same time as other reforms, such as standards-based teacher evaluations, so distinguishing the impact of PAR from that of other reforms proves difficult. Finally, many existing PAR programs are longstanding and began before states and districts collected comprehensive data on teachers and student outcomes. As a result of these challenges, there are currently no causal estimates of PAR's effects on important educational outcomes, although the flurry of recent interest in PAR presents a clear opportunity for future studies of the program's effects on teachers and students.

Second, recent work suggests that examining only short-run effects or specific, easily-measured outcomes may mask substantial longer-run benefits of educational interventions. Most educational program evaluations focus only on short-term student achievement test scores. However, evidence is mounting that interventions that show no immediate benefits and no increases in student test scores may substantially improve long-run outcomes, such as labor market earnings or participation in crime. For example, MDRC studied career academies in New York City schools using a randomized control trial (Kemple, 2008). They found that attending a career academy did not increase student test scores. However, in a follow-up study that tracked students eight years after graduation, they found that students who had won a lottery to attend a career academy earned 11 percent more in the labor market than students who entered the lottery

but were not selected. Thus, they concluded that students in these academies learned other skills that paid returns in the labor market. Similarly, Deming (2009) found that students who moved to better schools as a result of a school choice program in Charlotte-Mecklenburg, NC, had no better test scores, but were substantially less likely to commit crimes in the future. To the extent that such longer-term benefits are not counted in the impact evaluation, program effectiveness will be understated.

Third, in a traditional cost-benefit analysis, all estimates must be converted to monetary units. To do so requires making substantial, and sometimes unfounded, assumptions about the present discounted value of such things as improved test scores or greater educational attainments (Boardman et al., 2006). In his seminal 1983 primer on cost-benefit analysis, Levin notes this limitation, arguing that “it is not always possible to find appropriate ways to express the benefits of a project in monetary terms” (1983, p. 109). For example, the classic benefit-cost evaluation of the Perry Pre-School program (Barnett, 1996) estimated a substantial return on investment, but some key results were driven by assumptions about the cost of crime (e.g., that rape cost a victim \$65,229 based on a combination of direct losses, pain and suffering, and risk of death). The imprecision of estimates about both the intervention’s effect and the monetary value of benefits makes pure cost-benefit analyses of programs with wide-ranging effects virtually impossible.

These challenges led us to employ a hybrid approach in this study. In the following discussion, we first catalogue PAR’s financial costs and compare them to monetized estimates of the program’s short-term administrative cost-savings. We then describe PAR’s potential effect on student learning. Although it is difficult to quantify the consequences of having a more effective teacher, we present one attempt to do so. Finally, we examine broader and longer-term

organizational benefits, as defined by the program's stakeholders. This approach enables us to account for benefits that are inherently unquantifiable, but, of course, we must rely on stakeholder reports, which are inherently subjective. We present the views of a wide range of individuals – including administrators and union members – documenting the perceived additional benefits of PAR. Although these benefits cannot be quantified, they are important for policymakers and district officials to consider as they decide whether to adopt PAR.

Study Methods

The existing literature on PAR focuses on specific states or individual programs (e.g., van Lier, 2008; Koppich, 2004) or on broader discussions of PAR's effectiveness (e.g., Goldstein, 2007; Lieberman, 1998). In this paper, we use administrative and interview data from seven districts operating PAR programs in order to better understand the program's costs and benefits. We report here on one part of a larger qualitative study examining PAR programs. In the larger study, we sought to learn how these districts had implemented PAR and what challenges they faced, to understand how their PAR programs worked and the tradeoffs involved in different decisions about program design, and to solicit stakeholders' views about the effectiveness of PAR in their district. In other analyses, we focused specifically on the role of the consulting teacher, the labor-management relationship at the district level, and the effects of PAR on school principals (Author, 2009).

We selected a sample of seven districts that had adopted PAR (Toledo and Cincinnati, OH; Rochester and Syracuse, NY; Minneapolis, MN; San Juan, CA; and Montgomery County, MD). We identified a larger number of possible research sites through literature reviews, internet searches, and informal networking. In our research, we found some districts with programs that were called "Peer Assistance and Review," but did not include peer evaluations that could lead to

dismissals; they were essentially mentoring programs. As a result, we included only programs that integrated assistance with review that could lead to dismissal. We sought a diverse sample of districts, with a range of geographic locations, district financial resources, labor-management relationship histories, and sizes. In Table 1, we present a range of background characteristics about these districts and programs. We purposefully included some districts, such as Toledo and Cincinnati, with long-standing PAR programs and others, such as Syracuse and Montgomery County, with relatively new programs. Most districts included both a Novice program for teachers new to the district and an Intervention program for experienced, tenured teachers, but these components varied widely in size. However, most districts that have implemented PAR programs are urban school systems or large districts in inner-ring suburbs. Developing a PAR program in a smaller district is certainly possible, but would involve different logistical challenges, as well as different program costs and benefits, than those we discuss here.

INSERT TABLE 1 ABOUT HERE

Between December 2007 and April 2008, two researchers conducted site visits of 2 to 3 days in each district. Over the course of each visit, we interviewed approximately 25 individuals, including key union and district officials, members of the PAR Panel, current and former Consulting Teachers, and principals. In all districts, we interviewed the superintendent or associate superintendent and the teachers union president. We made a concerted effort to speak with both advocates and opponents of PAR. Across the seven districts, we interviewed 155 individuals, using semi-structured interview protocols, tailored to the specific sites and roles of those interviewed. The study's general interview protocols are available from the authors on request. We supplemented the interviews with reviews of relevant documents, including local collective bargaining agreements and program handbooks.

We transcribed all interviews and catalogued all documents for analysis. As a first step in data analysis, we wrote a thematic summary describing the program for each district. This allowed us to identify broad similarities and differences across the seven districts. Subsequently, we conducted detailed data analysis by coding documents and interview transcripts. We developed theoretical codes drawn from the research literature and open codes identified in both the thematic summaries and the interviews. We also created a number of matrices to identify patterns in the data, both within and across districts (Miles & Huberman, 1994).

We found notable agreement across all seven districts about the value of PAR programs and their perceived positive impact on teaching and learning. This is not surprising, since these districts had invested substantial resources in PAR and had continued to support the program over time. However, we did learn from others that, within some districts, the program was controversial when it was adopted because some teachers did not think that teachers should evaluate peers or that the union should sponsor a program that could lead to a teacher's dismissal. However, once the program had been well established, it was said to be widely endorsed by teachers. In selecting a small sample of principals to interview, we requested to talk with some who supported the program and some who opposed it. Consequently, in each district, we heard a range of opinions about PAR from principals. Here, too, initial opposition to the program subsided over time as principals saw that it enhanced rather than undermined their authority.

In each district and across the sample, we carefully examined the balance of comments about PAR, recognizing that our sample was small and purposive and, therefore, could not be used as the basis for generalization. However, deliberately gathering a range of views from union leaders, administrators, and consulting teachers enabled us to better understand individuals'

satisfactions and dissatisfactions with PAR. We have included all perspectives in our discussion below. However, on the particular topic of costs and benefits, opinions were less divided than they were on several other topics, such as whether teachers should be given sole responsibility for peer evaluations. In explaining what we learned, we rely primarily on quotations, descriptions and examples. When appropriate, we use words such as “many” or “most” to convey the prevailing views or experiences in the large majority of districts or among sub-groups of respondents, such as principals.

As we explain, the stakeholders we interviewed said that PAR’s benefits far outweigh the program’s costs. That assessment is not surprising, given that these districts adopted the program and have chosen to continue it. Other districts, which we did not study, may have had less successful experiences with PAR and, therefore, may view benefits and costs differently. However, the broad-based support for PAR – particularly among central office staff – in a wide range of districts, including those that had recently adopted PAR, suggests that it has great potential for success in many locations.

Section III. The Financial Costs and Short-Term Benefits of PAR

Financial Costs

To enumerate the financial costs of PAR, we employed the “ingredients method” set forth by Levin (1983) in which we valued each of the inputs to PAR programs. Our analysis, though, was neither simple nor straightforward. Many districts funded PAR as part of a larger program

and administrators had difficulty separating out the resources associated only with PAR.³

Nonetheless, our analysis of district expenditures reveals several key patterns.

Although the program costs depend on the extent of PAR services provided, this is a relatively expensive reform. The districts we studied exhibited wide variation in their PAR expenditures, reflecting variation in program size and complexity. Larger programs and those with more components cost more. For example, it takes more resources to serve struggling veterans who risk dismissal than new teachers, many of whom are likely to succeed. Because of this, districts adjusted CTs' caseloads so that a teacher on Intervention counted the same as several novices. We take these differences into account when we evaluate the cost per teacher in PAR, which ranged from approximately \$3,000 to \$7,000 for novices and \$6,000 to \$10,000 for teachers on Intervention.⁴ In our analyses, we use an average figure of \$4,833 per novice teacher and \$8,350 per teacher on intervention.

In Table 2, we present a detailed accounting of the program costs from Toledo, a typical program and one that has maintained detailed data on program costs. In Toledo, total program costs approached \$800,000 to serve 124 novice teachers and 16 experienced teachers. Here, and in all districts with PAR, by far the greatest expense comes from hiring teachers for the classes of CTs who are released either full-time or part-time. Although these substitute teachers generally have less experience and earn less than the CTs, the total cost of their salary and fringe benefits is significant. Across districts, replacement expenses represented approximately 75% of PAR expenditures.

INSERT TABLE 2 ABOUT HERE

³ For example, in Cincinnati the entire district uses peer evaluators as part of its Teacher Evaluation System. Thus, the same individuals serve as CTs in the PAR program and as Teacher Evaluators for other teachers not in PAR. Accounting for the portion of their salaries and benefits devoted solely to PAR proves challenging.

⁴ Districts typically count each Intervention teacher in a CT's caseload as a more than one "case". We adjust the total resource costs by the district's weighting of novice and Intervention teachers in caseload determination.

Although all other expenses pale in comparison to these replacement salaries, districts incur a wide range of additional costs. CTs earn a stipend, ranging from \$5,000 to \$10,000 across the programs studied, and some districts also provide stipends for PAR Panel members. Districts need to train CTs and Panel members, although several use current CTs to provide the bulk of the training as part of their roles. Because Panel meetings are typically held during the school day, teachers and principals on the Panel may need substitutes to cover for them during meetings. While the PAR Panel oversees the program, districts also need some way to facilitate day-to-day program management. Some distribute these responsibilities across several “lead” CTs, but others, particularly those with more extensive or complicated programs, hire a full-time program director. Finally, districts must cover the program’s administrative costs, including office space, computers, mileage reimbursements for CT travel, and administrative/clerical support staff. Several districts have found creative ways to limit these expenses; for example, the Syracuse teachers union provided office space in kind at no financial cost to the district, while other districts housed their CTs in closed school buildings. Whatever their strategy, districts must find some way to cover these administrative costs.

With the exception of a few one-time expenses, the primary start-up costs of implementing PAR are no different than the ongoing costs. Although established models for PAR programs are available, districts must invest some resources in planning for, adopting, and implementing an effective program in any district. PAR is not simply an add-on to the district’s current evaluation system. Rather, PAR must adjust to current programs and these other initiatives must adjust to accommodate PAR. However, these up-front investments pale in comparison to the costs of hiring replacement teachers in a continuing PAR program.

Accounting for Existing Programs

Importantly, the per-teacher costs presented above reflect the full program costs, not the net or marginal costs after accounting for other programs that PAR replaces. However, PAR often does take the place of traditional evaluation and mentoring of first-year teachers, which most districts already provide at some expense. Even relatively inexpensive mentoring programs provide a stipend for mentor teachers and require some administrative costs. Intensive mentoring, such as that sponsored by the New Teacher Center, can cost even more than PAR, between \$6,000 and \$7,000 per teacher (Villar & Strong, 2007). In their study of a mid-western school district, Milanowski and Odden (2007) estimated that school-based induction costs \$4,525 per new teacher. PAR would represent only a very moderate cost (\$300) above the current induction practices in this district. As a result, enumerating the total financial costs of PAR to a district remains quite difficult because it depends in large part on the district's current expenditures, particularly in programming for novice teachers.

Administrative Cost-Savings

Two main benefits of PAR have clear financial implications for districts: increased teacher retention and lower arbitration and dismissal costs. Both of these represent substantial administrative costs for district human resources departments. Our analysis suggests that PAR's Intervention program saves districts money and is cost-effective simply based on short-term administrative cost-savings. The novice program, on the other hand, would need to produce an enormous reduction in turnover to justify its costs on this basis alone.

Evaluating the short-term administrative savings is difficult because costs of turnover or dismissals rarely show up as line-items on district budgets. Still, they represent large potential expenditures for most districts. The cost savings from PAR vary widely from district to district and depend largely on the district's current practices. As a result, we do not present a "bottom-

line” figure of savings to compare to program costs. Instead, we provide an assessment of the types of improvements that would be required to justify PAR based on these administrative cost-savings. Again, given that our cost estimates do not reflect marginal costs over existing programs, we must consider these projections as relative improvements compared to a district with no induction program.

Reduced costs of turnover

Although there is no definitive causal evidence that PAR reduces turnover, districts credit PAR with increasing retention substantially. For example, Rochester’s first-year teacher retention rate increased from 60% to nearly 90% after introducing PAR (National Commission on Teaching and America’s Future, 1996). In fact, in all the districts we studied that reported turnover data, first-year teacher retention averaged approximately 90%, which far exceeds the national average for urban districts. The Rochester Superintendent attributed the district’s low turnover specifically to PAR: “A little bit over 70 percent of my budget is spent on people: \$680 million. It’s all about people, it’s all about talent. And [PAR] is a great process in developing great teachers and retaining great teachers. The retention is amazing.”

Other district officials pointed to the program as a highly successful component in their efforts to retain new teachers. In Syracuse, a principal noted that PAR led teachers to remain in the district because they became more effective in the classroom: “Teachers were quitting and jumping to the suburbs. But, I think with PAR they are really learning how to teach urban kids and they are really excited about it because their skills are so well-developed so early on in their career. They are committed in wanting to stay.” This sense of effectiveness is an important determinant in turnover decisions of early career teachers (Johnson & Birkeland, 2003). PAR can also affect teacher retention well beyond the early stages of a teacher’s career by raising the

performance of veteran teachers who enter the Intervention program and by offering differentiated roles outside of the classroom for expert teachers, thus enriching their careers.

PAR's potential to improve teacher retention is important because studies suggest that the financial costs of teacher turnover are high, since districts must expend central office and local school resources to recruit, hire, and train new teachers. Estimates of the actual costs of turnover range widely by school district and by assumptions used to determine these costs, but in all cases they are quite high. For example, the Texas Center for Educational Research (2000) used different models to produce estimates that ranged from \$6,060 to \$48,480 per first-year teacher replaced. Other, more data-driven estimates suggest that urban districts spend approximately \$10,000 to \$20,000 to replace a novice teacher who leaves, and substantially more to replace teachers who leave after longer careers because of the experience, knowledge, and training that they have accumulated (Barnes, Crowe, & Schaefer, 2007; Birkeland & Curtis, 2006). Using data from the Boston Public Schools, Birkeland and Curtis (2006) estimate that, in 2004-05, it cost the district \$10,547 to replace a novice teacher. Milanowski and Odden (2007) estimate replacement costs of \$9,061 for novice teachers in a large midwestern district. Barnes, Crowe, and Schaefer's (2007) estimates of turnover costs are \$17,872 in Chicago and \$15,325 in Milwaukee.⁵ Importantly, these estimates include only administrative costs, not the costs of lost productivity and organizational instability that result from turnover.

Assuming, conservatively, that a district incurs costs of \$10,000 when it replaces a novice teacher, PAR would have to reduce first-year teacher turnover by approximately 48 percentage points to offset the full program costs with the benefits gained from increased retention. It seems unlikely that PAR would have such an enormous effect. However, to the extent that PAR's

⁵ These figures are not for novice teachers. However, Milwaukee only included \$129.01 for training of all teachers in their cost estimates, suggesting that the \$15,325 figure closely matches the replacement cost of a novice teacher.

marginal costs over an existing program are less than our estimated \$4,833 per first-year teacher, the program could offset its costs with a less dramatic improvement in retention rates.

Furthermore, \$10,000 is a conservative estimate of the costs to replace a new teacher, and replacing more experienced teachers imposes even greater expense. For example, Birkeland and Curtis estimate that it cost Boston Public Schools \$26,687 in administrative, hiring, and lost professional development costs to replace a third-year teacher. This estimate, appropriately discounted, implies that a reduction in third-year teacher turnover of 9.2 percentage points would fully offset PAR program costs. Nonetheless, although the estimates are sensitive to a variety of assumptions, it appears that the administrative cost-savings from reduced turnover alone cannot fully justify PAR's expenses.

Reduced costs of teacher dismissal

The district officials we interviewed argued that PAR does not simply reduce overall turnover, but also promotes selective retention. The program helps to retain teachers who are succeeding, identify those who need help, and move out those who fail to meet standards. Thus, PAR can help districts improve their personnel practices and avoid the costly mistake of granting tenure to teachers who do not meet the district standards. The superintendent in Syracuse argued that there are “immeasurable” costs involved with rehiring and eventually offering tenure to “a first-year teacher that isn't great.” A central office administrator in Minneapolis concurred: “There's a huge benefit [to PAR] because, I think number one, if you tenure somebody that's not ready, those are million dollar decisions.” In part, this cost of offering tenure to an under-performing teacher plays out in future dismissal costs.⁶ Therefore, PAR can represent

⁶ In many cases, dismissal may not even be pursued. If not, the costs are then borne by students who must endure poor teaching; such costs cannot be quantified but are clearly greater than the costs of pursuing a dismissal.

substantial cost-savings by preventing such teachers from achieving tenure and reducing the costs of dismissing tenured teachers who do not meet district standards.

Dismissing teachers with tenure is ordinarily a very expensive and time-consuming process, which very few districts actively pursue. There has been no national study that quantifies dismissal costs; however, the legal fees, arbitration costs, and the expenses associated with several levels of appeals create substantial expenditures. Bridges (1992) interviewed district administrators in the 1980s and found, even then, that districts expected to pay \$50,000 for a single dismissal, with some proceedings costing upwards of \$200,000. A recent analysis by the New York State School Boards Association found that the average cost of legal fees from pursuing a disciplinary case in New York was \$128,941 (Honawar, 2007). In addition, these figures do not count the substantial amount of time that principals and district Human Resources administrators must spend to document and build the case.

The districts we studied all have succeeded in removing tenured teachers without incurring expensive legal challenges (see Table 3). Although the districts formally dismissed no more than ten teachers per year, these dismissal rates far exceed those before PAR. Furthermore, tenured teachers assigned to the PAR Intervention program often choose to resign or retire instead of facing formal dismissal. Such teachers were said to see “the writing on the wall,” realizing that both the union leaders and administrators supported their removal. In Rochester, the union president described how he talked with underperforming teachers: “the union, without batting an eye, says, ‘If you’re not doing the kids any good and we can’t help you come up to that level, then you’re not doing the rest of us any good.’” Because such frank conversations often lead veteran teachers to resign or retire, it is not enough to simply count formal dismissals in assessing PAR’s impact on the district’s teaching force.

INSERT TABLE 3 ABOUT HERE

When districts do proceed to formal dismissal proceedings, the detailed PAR procedures and thorough review by CTs ensure that teachers' due process rights are met and that union leaders meet their legal duty of fair representation. Because they are not obliged to support the appeals of teachers dismissed under PAR, the union leaders can maintain high professional standards for teachers. The Syracuse union president expressed this attitude: "Frankly, I think that there is a difference between representing people and defending the indefensible. I think in this day and age, in this age of accountability, true accountability, we have to draw a line." Every union official we interviewed refused to pursue appeals for teachers who had reached the end of the PAR process.

As a result, dismissal proceedings that do occur move smoothly. A Syracuse district administrator saw the real financial advantage of having labor-management collaboration in dismissal cases: "We're saving money because, with PAR, we're not going for arbitration with the union and all those other hearings, because the union is sitting at the table from the beginning. The union is listening to each presentation that is being done by a union member as a consulting teacher." An administrator in San Juan concurred, saying, "The union is certainly with us every step along the way." These district officials were convinced that the joint involvement of union members throughout the PAR process saved the district substantial legal costs in processing dismissals.

Moreover, the teachers themselves are less likely to appeal decisions on their own because the recommendations come from fellow teachers through a detailed, evidence-based evaluation process. According to a union representative in Minneapolis, "It's just very difficult to contest [a negative judgment] when you've been told by your peers that you're not

performing.” Thus, because PAR is a comprehensive process built on strong labor-management collaboration, districts reported that very few dismissal decisions were challenged. As a result, the San Juan Superintendent said, “I mean, yes, if you can ... avoid the costs of some dismissals, it doesn’t take very many to pay for a pretty extensive program!”

Although PAR Intervention programs, at an average cost of nearly \$8,400 per teacher, are substantially more expensive than the Novice components, the exorbitant costs of teacher dismissal proceedings clearly justify the program’s expenses for any district that would otherwise pursue dismissal. Across the districts we studied, approximately two-thirds of teachers on intervention failed to successfully complete the process and left the district. Given the costs of dismissal, the program would be cost-effective if only one out of every fifteen teachers on Intervention were dismissed or chose to resign.⁷ In districts such as Montgomery County that dismiss, on average, five tenured teachers a year, PAR’s Intervention program pays for itself many times over simply in administrative cost-savings. Notably, this analysis does not account for the improvements in student achievement and productivity that result from dismissing persistently underperforming teachers.

IV. PAR and Student Achievement

In the previous section, we focused on the financial costs and administrative cost-savings from PAR. However, above all, PAR is a program designed not to save districts money but to improve teacher quality and instruction. Thus, PAR’s effects on student outcomes, including student achievement, must be factored in to any cost-benefit analysis. Although the extent to which PAR improves student achievement gets at the heart of what policymakers value, there

⁷ Using an average cost of teacher dismissal case of \$128,941 and the average cost of PAR Intervention programs of \$8,350 per teacher.

have been no quantitative evaluations of PAR's effects. Nonetheless, the most common benefit cited by our study participants was improved teacher quality; interviewees recognized PAR as a key component in their efforts to ensure that every student would have an effective and competent teacher. It is difficult to quantify these benefits, particularly the effect on students of having an ineffective teacher. Here, we present evidence that PAR likely raises student performance and offer one preliminary attempt to quantify some of the monetary benefits that may result.

PAR's Effects on Teacher Quality

PAR operates through several avenues to raise student achievement. First, the Novice component provides a strong induction program. Key stakeholders in these districts said that PAR helps new teachers get off to a successful start under the watchful eye of an engaged and skillful mentor. In PAR, an experienced CT observes the teacher almost weekly, provides detailed feedback, both formally and informally, and helps support the teacher in her professional growth. As a result, one Minneapolis CT said, "these new teachers soar." Furthermore, as we described above, PAR appears to improve teacher retention. Educational researchers agree that more experienced teachers outperform their novice colleagues in the classroom (Rockoff, 2004; Clotfelter, Ladd, & Vigdor, 2006). As a result, retaining more teachers who have been effective in their first year will contribute more to student achievement than replacing them with other novices.

Similarly, the Intervention program helps many underperforming experienced teachers get back on track. Teachers on Intervention may be struggling in a new teaching assignment, be dealing with personal problems, or simply have lost the motivation to improve their practice. With intensive mentoring, many teachers improve and go on to succeed, while those who cannot

meet district standards are dismissed. A PAR Panel member in San Juan explained this dual role of Intervention: “We’ve saved a lot of careers. Those teachers are now in the classroom doing really good things. [At the same time], we’ve gotten some people that needed to be moved out, and we got them out quicker than in the old process.” Across the districts, CTs, principals, union officials, and district administrators acknowledged that PAR had provided just the spark some teachers needed to revisit their practice with expert assistance and to return to their classrooms as effective teachers. And, in cases where teachers did not improve, PAR provided an effective route for termination, either through informal counseling out of the profession or formal dismissal.

For teachers in both the Novice and Intervention programs, CTs’ work was widely seen as superior to the traditional evaluations done by principals. Because working to support and evaluate teachers is their full-time job, CTs can spend much more time with individuals on their caseload than can principals, who must manage all aspects of the school. According to a Toledo principal, “if you are in a position to have a consultant who can go in weekly, observe these people, and make arrangements for them to go out and observe other teachers, I think that is of great benefit.” By visiting a teacher’s classes weekly rather than dropping by intermittently as most principals do, a CT can offer more targeted feedback to help the teacher improve and can assess the teacher’s potential, needs, and growth over the year.

Many also said that the CTs could provide more objective evaluations than principals. A central office administrator in Cincinnati argued that this objectivity comes from the CT’s position as an outsider in the school:

This is a person who comes and does the observation, who has nothing to gain one way or the other. If you’re outstanding they have nothing to gain, and if you’re lousy they have nothing to gain. So it isn’t, “Well, this is the principal who is just angry with me. She’s mad because I didn’t organize the Christmas party last year.” It’s very, very

objective.

A Syracuse principal agreed, saying that she tends to give her novice teachers the benefit of the doubt, even when they do not deserve it, “because I have established a bond with that person.” Thus, she saw the CT’s assessments as more objective. Some also said that CTs’ advice and judgment carried greater weight with teachers in the program because the CTs were known to be effective teachers. A former principal in Minneapolis saw this peer relationship as key to PAR’s success: “In my role as a principal, I can say exactly the same thing that a teacher colleague can say, and because it came out of my mouth and not the teacher colleague's, it’s interpreted differently.”

Across the districts, then, administrators and union leaders thought that PAR was an important investment because it improved instruction. A Cincinnati administrator said that PAR “does what many evaluation systems don’t do, and they’re not designed to do, which is to improve teaching performance.” In particular, many stakeholders talked in detail about the large price that students pay when they have an ineffective teacher. A PAR Panel member in San Juan referred to the “credit card commercial” in talking about PAR, saying that the program “is priceless because it impacted 20 little kids for the rest of their life to have a quality teacher.” This sentiment was echoed by teachers and administrators across all districts in our study.

Several stakeholders, such as the Cincinnati teachers union president, did note the concern that PAR removes excellent teachers from the classroom: “You have got twenty teachers who are arguably supposed to be your very best of your best out of the classroom, not teaching kids, but evaluating other teachers.” However, many stakeholders directly rejected this criticism. One Syracuse principal, who summarily dismissed that claim as “a bunch of baloney,” argued that the CTs “are impacting the students of all of the interns and ... they’re leaving [the

classroom] for three years, and then they're going back. So, knowing that, that's the best of both worlds, because you're leaving to impact many more students." That CTs could have a broader, albeit indirect, influence on many students across the district convinced most stakeholders that the costs of removing a small number of expert teachers were easily counterbalanced by the improvements in other teachers' work.

Quantifying PAR's Effects on Student Achievement

While stakeholders were convinced that PAR improved teacher quality and raised student achievement in their districts, quantifying these benefits proves to be quite difficult. We attempt to do so by examining only one part of PAR's overall effect – the improvement in teacher retention. That we do not include additional benefits – such as the removal of an underperforming veteran teacher or the benefits of having a strong induction program – suggests that our estimates here understate PAR's true value. We find that even very small improvements in student achievement from PAR would be sufficient to justify the program's costs.

Even lower bound estimates of the returns to teacher experience suggest that first-year teachers are much less effective than second-year teachers in raising student test scores, by approximately 0.05 standard deviations (Rockoff, 2004; Ladd et al., 2004; Ladd et al., 2009).⁸ Valuing this improvement in student achievement is tenuous, but even very conservative estimates suggest that the simple labor market returns to this improvement would be substantial. Using Murnane, Willett, & Levy's (1995) estimates of the return to cognitive skills from the High School and Beyond survey, we estimate that a 0.05 standard deviation improvement in test scores would raise annual earnings by 0.40% for men and 0.55% for women. Average annual earnings in 2008 were \$48,678 for men and \$29,609 for women (US Census, 2009). Assuming

⁸ We term these estimates "lower bounds" because they only account for the performance of students in the teachers' classroom and do not account for positive externalities on peer teachers (see Jackson & Bruegmann, 2009) or other students in the school through better institutional coherence.

that these estimates of the return to skills are valid for students in these districts and that real wages are constant, these returns translate to an annual wage increase of \$193 for men and \$163 for women. Appropriately discounted and with a range of very conservative assumptions, such as a 5% real discount rate,⁹ these returns translate to a present-value benefit of \$192,962 per novice teacher who remains in the district for only one additional year.¹⁰ Obviously, if PAR induces a novice teacher to remain in the district for more than one year, this estimated return would more than double.

Adding these productivity returns to the administrative costs clearly changes the calculation. Here, PAR would need to improve novice teacher retention by only 2.3 percentage points – a very modest goal – to justify its costs. In other words, in a district with 50 new teachers in PAR, the program would need to encourage only one additional teacher to stay in the classroom to make it cost-effective. Again, this analysis focuses only on retention for first-year teachers and ignores all of the other avenues through which PAR may affect student achievement. The bottom line is that a teacher-level intervention like PAR only needs to improve student performance very minimally to satisfy a cost-benefit test.

V. Organizational Benefits of PAR

⁹ These estimates are conservative in several ways. First, Murnane, Willett, and Levy's (1995) estimates derive from equations that include educational attainment as a covariate. As a result, they understate the true return to improved academic achievement because better cognitive skills lead to greater educational attainment (we thank Richard Murnane for pointing this out). Second, Murnane, Willett, and Levy found that the returns to cognitive skills had increased substantially between 1972 and 1988. Given the growing value of education in the labor market, the return to skills is likely greater today than in 1988 (Katz & Goldin, 2009). Third, we assume a real discount rate of 5%, which is very conservative. Finally, these benefits account only for labor market returns and not for improvements in health care, crime, or a range of other outcomes that cost society substantial amounts of money (see Belfield & Levin, 2007). It is important to note that these estimates are based on a partial equilibrium model, which is appropriate given the small scope of PAR programs.

¹⁰ We assume that students enter Kindergarten at age 5, begin work at age 20, and retire at age 65. We also assume that each elementary school teacher teaches 20 students and each middle and high school teacher teaches 80, evenly divided between men and women. Finally, we assume that teacher turnover is evenly distributed across the grades.

In the previous analyses, we attempted to quantify the benefits of PAR in terms of both administrative cost savings and improvements in student achievement resulting from increased teacher retention. These estimates, however, do not capture the full range of benefits that districts gain from PAR. Although several district officials acknowledged that PAR could be justified purely on the basis of a financial benefit-cost analysis because of improved retention and reduced dismissal costs, very few framed their responses in financial terms. Instead, district officials, union leaders, principals, and consulting teachers all talked much more expansively about the benefits of PAR for teaching and learning. As a central office administrator in Rochester asserted, “the cost doesn’t bother me at all. ... [It] can be very, very expensive. But I think it’s an essential investment in launching teachers in our district successfully.” As seen in Table 4, union and district officials saw PAR as an investment in human capital that had payoff in four main areas, all of which related indirectly to improved student performance: building a positive organizational culture centered on instruction, improving labor-management relations at the district and school levels, alleviating burdens on principals, and providing professional development and career opportunities for consulting teachers.

INSERT TABLE 4 ABOUT HERE

Building a positive organizational culture centered on instruction

Many stakeholders said that PAR can contribute to building a strong, professional culture of teaching. Through PAR, expert teachers can become the gatekeepers to the profession and help to uphold and enforce professional standards. By taking the stand that underperforming teachers who cannot meet district standards should be removed from their positions, union leaders contradict the conventional view that unions protect all teachers at all costs. This attitude has lasting repercussions in the district; in Syracuse, a principal said that “the joint development

of this program between the district administration and the teachers union really shows a commitment to professionalism.”

PAR also helps to challenge and reshape longstanding norms of autonomy and privacy in public schools. As new teachers come to expect that CTs will be present and involved in their classes regularly, the deeply entrenched norms of the egg-crate school begin to erode. Thus, PAR can encourage collegiality and open practice. The union president in Minneapolis saw this cultural shift as a key benefit of PAR: “I think you have to factor in the elements that you can’t quantify in terms of dollars, just the – sort of – general climate that I think is going to be more positive when you have this constructive assistance.” The Syracuse Superintendent agreed, saying that “to go over all the litanies of failures in the [traditional] evaluation process is probably pointless, but clearly until we create an atmosphere where teachers are in a culture of change, in a culture of support, we’re not going to get the kind of instruction that we want to happen.” He said that PAR promotes such a culture in Syracuse. As more and more teachers go through PAR, these cultural changes can extend throughout the district.

PAR’s steady focus on teacher evaluation was also said to stimulate dialogue about effective instructional practice. Teachers who have been evaluated in PAR often know the district standards very well. Several individuals described their district’s renewed focus on discussions of practice, and they attributed that emphasis to PAR. In fact, a Cincinnati program administrator said that PAR’s “biggest benefit is getting this dialogue, continuing this dialogue about what good teaching practice is.”

Improving labor-management relations at the district and school levels

In addition to encouraging collegiality among teachers, PAR can also strengthen a collaborative labor-management relationship, not only saving the district money as fewer

grievances move to arbitration, but also working to build a positive approach to reform throughout the district. A Syracuse principal described these benefits: “it does something for the relationship between the union and central office administration, because you really feel like you’re working together for improved things.” Principals across the seven districts saw this collaboration trickling down to the building level, improving relationships between union building representatives, other teachers, and administrators in the school.

In large part, the regular participation of teachers union officials and administrators on the PAR Panel helped build a relationship of trust and collaboration. As one Panel member in Syracuse said, “in the beginning, [the Panel] wasn’t an adversarial group, but we were all kind of put together and these are your roles and here were the administrators and here we [the teachers] were.” Over time, though, “it became like a common purpose together.” The Panel chairs, typically high-ranking union and administration officials, work together even more regularly. For example, the union co-chair in Rochester said that her administrative counterpart “happens to be in a position where I speak to her almost daily. She is the chief human resources officer, and shoot, I talked to her twice I think today. She was over here for a meeting and then I called her later on.” This type of regular communication not only made PAR possible, but also facilitated mutual respect and collaboration when other labor-management issues arose.

Improved labor-management relationships can become important throughout the district. Traditionally, the relationships between teachers unions and district administrators have been adversarial. Bargaining typically involves each side stating positions and then bartering over specific provisions to be included in any agreement, while day-to-day relationships center on ensuring compliance with negotiated agreements. Improved relationships through PAR can help lead to more collaboration and an increased focus on student learning.

Alleviating burdens on principals

Some principals in the districts that we studied saw PAR as an encroachment on their rights and responsibilities as building administrators. However, the large majority appreciated the program. This was particularly true in districts with longer-standing programs. By providing steady support and evaluations for novices and underperforming veterans, CTs reduce the burden on principals to evaluate all teachers. Importantly, though, principals still have many opportunities to assess teachers in their second year and beyond.

Across our study sites, principals reminded us that running a school is a complex and difficult task, requiring attention to many demands and details, often pulling them away from the teachers who most need their support—novices and underperforming veteran teachers. A Minneapolis principal echoed a common theme heard from principals in many districts: “As a principal of a large school – we have a thousand kids here – I wouldn’t have that time to devote to a new teacher as well as they do. So, I have had tremendous results from the program.” Similarly, a consulting teacher in Rochester recalled a principal saying, “I’ve got so much on my plate that I’m not going to have a whole lot of time to visit with this intern. So, I’m counting on you to support and help this young man along.” Principals recognized the added value that CTs brought to their schools, including specialized expertise in content areas. No principal, however hardworking, could adequately advise teachers about instruction in every subject. Furthermore, given that principals earn substantially more than teachers, a CT’s time is less costly than a principal’s time.

Beyond the time savings and improved evaluations that result from PAR, many principals also appreciated that the program enabled them to take on a more supportive role, particularly by allowing them to build more personal connections with their novice teachers. As one principal in

Syracuse explained, “I am able to relax in that sense and know that the PAR teacher is coming in to do some really good things. And, I can watch that teacher from a different point of view, which is really important to me because I am able to see their development more.” A Cincinnati principal agreed, saying “I actually think it’s a support to principals that we have someone who is qualified to come in and model lessons and demonstrations and provide the extra support and assistance to this teacher who desperately needs it.” He liked that he could identify a problem but then rely on the CT to provide the necessary mentoring to help the teacher improve.

Providing opportunities and professional development for consulting teachers

Finally, PAR benefits the consulting teachers and, in doing so, may enhance the opportunities and attractiveness of a career of teaching. One common feature of the teaching profession is its “flat” career path, with thirty-year veterans often doing the exact same job as novices. Serving as a PAR CT offers a new role and additional responsibilities for experienced teachers who are highly skilled. Many of the current and former CTs we interviewed spoke of PAR as contributing to their learning and development as professionals. In Toledo, for example, one teacher called her time as a PAR CT “the biggest, most important professional growth of my entire career.... It changed me forever.” In districts, such as Toledo, that have no differentiated career structure for teachers, PAR provides opportunities that, for many, encourage them to remain in teaching. For districts, such as Montgomery County, that sponsor a selection of specialized roles for teachers, PAR provides valuable experience that often leads to further career opportunities.

Current and former CTs consistently reported that the role offered professional challenges that rejuvenated them. A Toledo CT said that “it’s been a renewing experience for me.” Most CTs talked about being excited to return to the classroom because they had been

inspired by a sense of new possibility. For example, one spoke of keeping a file on her computer with new things she wanted to try when she returned to the classroom. According to one Cincinnati CT, “I really enjoy going into classrooms and seeing fantastic teaching, because I am stealing every good idea that’s out there.”

Beyond the value of seeing promising new teachers in action, CTs also said that their role as evaluators helped them develop a better understanding of what makes for good and successful instruction. In Rochester, a union official said, “A persistent recurring comment is, ‘I know my interns benefited but not as much as I did. I learned more about teaching this year than I did in my first 20 years in the classroom.’” The Syracuse Superintendent saw this as one of the most important benefits of PAR:

These six teachers that are PAR consultants are going to be fabulous when they go back to the classroom. I mean, they are already good. They will be incredible and if we keep turning them over, you’re creating a core of just incredible educators, because I’m convinced that you don’t know good teaching until you see somebody else teach. We all do this in isolation and there’s no possible way that we can really perfect our skills.

Former CTs regularly acknowledged that their clear understanding of the district’s instructional standards stayed with them when they returned to the classroom.

CTs not only reported that they had acquired new techniques to enrich their own teaching practice, but also said that they had returned to the classroom with a new perspective on the district and its challenges, which enabled them to exercise informal leadership within their schools. One explained, “I have some clout. And I have some understanding of the system. And I can make things happen at my school that maybe I wouldn’t have even dared to try before, didn’t know how, didn’t have the courage, didn’t have the confidence.” Others said they had gained the confidence needed to take on a new challenge, either a new subject area or an assignment in a more challenging school.

Although most resumed teaching full-time, other CTs moved into formal school-based roles upon completing their term, serving as instructional coaches or curriculum specialists. New doors had opened for them because of their experience as CTs. Although most districts discouraged CTs from assuming administrative roles for at least a year after their term ended, some CTs reported planning to become principals as a result of their work in PAR. Thus, within districts, PAR served to strengthen the human capital system by developing both teacher leaders and future administrators. For new teachers, who at first may not have planned a long-term career in teaching, PAR could encourage and extend their commitment by providing the prospect of new, professional opportunities.

Section V. Conclusion and Implications

In this article, we have developed and then applied a multi-staged approach for evaluating the costs and benefits of an educational program. For many reasons, a straightforward cost-benefit or cost-effectiveness analysis of PAR remains difficult, or impossible, to complete. As a result, our approach consists of three components. We first account for the well-defined financial costs of PAR and examine the magnitude of short-term administrative cost-savings that would be required to justify the program. PAR is an expensive program, with costs ranging from approximately \$3,000 to \$7,000 per novice teachers and \$6,000 to \$10,000 per teacher on intervention. These costs are largely driven by the salaries and benefits of teachers who replace the CTs in their classrooms. These total costs, however, overstate the actual financial impact of PAR because the program often replaces other expensive programs, including mentoring and professional development for novice teachers.

In many ways, PAR can save districts money simply by reducing administrative costs from teacher turnover and dismissals. We find that, on administrative cost-savings alone, the PAR Novice program may not be cost-effective because PAR would need to reduce teacher turnover substantially to offset program costs. Given the large costs of pursuing teacher dismissals, though, PAR's Intervention program is cost-effective based entirely on short-term administrative cost-savings.

Second, we turn to the real intent of PAR – to improve student outcomes. PAR improves teacher quality and classroom instruction in several ways, by increasing the frequency and effectiveness of classroom evaluations, helping districts retain effective teachers, and encouraging districts to dismiss teachers who are not succeeding with their students. In the future, with more rigorous evaluations of PAR, a more precise accounting of these short-term benefits on student achievement may be possible. Nonetheless, we find that any minor improvement in student achievement would be more than enough to justify PAR's costs. For example, if we simply count the loss of productivity produced by teacher turnover in the first year, we find that PAR's novice program would only need to improve first-year teacher retention by 2 percentage points to be cost-effective.

Finally, we examine the organizational outcomes of PAR because a quantitative cost-benefit analysis of PAR does not fully account for the broader, longer-term benefits of PAR. The stakeholders we interviewed tended to consider PAR's costs and benefits in these terms. They argued that PAR encouraged a culture of collaboration not only among teachers, but also between labor and management at both the school and district levels. It focused teachers' attention on instruction and promoted conversations about sound practice and instructional

standards across the district. Finally, PAR had positive effects on the roles of principals and consulting teachers.

Thus, on balance, the individuals with whom we spoke saw PAR as an expensive program, but one that they judged to be worth the cost. The Toledo union president said, “It is the best investment the district makes, all about teacher quality for every reason.” Similarly, according to the Syracuse Superintendent, “This is a huge investment for us and even in a time when we cut sixty-seven positions last year, we didn’t cut this.” That superintendents, central office administrators, and union leaders, including officials in districts facing budgetary shortfalls, supported PAR so vocally lends credence to the conclusion that the program can have real and lasting benefits on schools.

Although there are to date no rigorous program evaluations of PAR, policymakers must often make decisions with limited information. Our analysis suggests not only that PAR can represent a key component of a district’s human capital strategy, but also that this investment, albeit considerable, may actually save districts money. Factoring in the non-financial benefits, including improved teacher quality, tips the scales even further in the direction of PAR. Although replacing existing induction programs with PAR might seem to represent an expensive option, choosing to invest in PAR may yield both financial and organizational benefits, which make it a smart use of funds, particularly in this time of financial constraints. Implementing PAR is a challenging undertaking, requiring clear commitment from both union leaders and district officials. However, our analysis suggests that districts that have chosen to implement PAR have invested up front in a system that can provide not only short-term cost savings but also longer-term payoffs.

Implications for Studies of Educational Costs and Benefits

Although we have focused on Peer Assistance and Review, our study suggests, more generally, several key lessons for cost analyses of educational programs and reforms. First, analysts must understand thoroughly not only the program itself but also the types of efforts that it replaces. Given that educational reforms typically draw resources from other programs, the analysis must account for the marginal costs and benefits of the new program. Thus, in large part, weighing the financial impact of implementing a program is contingent on the district's current practices and investments.

Second, our analysis suggests both that rough cost-benefit analyses are possible and worthwhile for assessing educational interventions, even without clear quantitative evidence of the program's outcomes. Educational policymakers and district officials must continue to decide which programs to fund, even when they lack complete and accurate data about costs and benefits. They cannot wait until rigorous methods are developed that allow them to quantify organizational costs and benefits, and these costs and benefits should not be ignored. At the same time, we find clear cause to advocate for more comprehensive evaluations that provide better estimates of a program's effects. These evaluations should not only be methodologically rigorous but also examine the full range of benefits that the program may yield. For PAR, recent adoption of the program in many districts provides promising opportunities for rigorous impact evaluations.

Finally, and most importantly, our study highlights the value of systematically approaching cost analyses of educational reforms. It is clearly important to capture accurately the relevant financial data, to account clearly for all of the relevant costs, and to understand the benefits of these reforms in financial terms. However, analysts should not limit their scope to the elements that are financial, or even those that are quantifiable. Instead, a comprehensive cost

analysis should also assess both the organizational program costs and the potential organizational benefits. The fact that we cannot quantify these organizational outcomes does not mean that they do not exist. Nor should they be ignored as policymakers, union leaders, and district officials decide whether to invest in an educational reform. We have laid out these organizational considerations by presenting evidence from program stakeholders about PAR. We recommend this more comprehensive approach for evaluating other educational programs that may have wide-ranging effects which cannot be easily quantified. Particularly in institutions as complicated as schools, examining these organizational issues fully is an important complement to a rigorous cost analysis.

References

- Author. (2010).
- Author. (2009).
- Barnes, G., Crowe, E., & Schaefer, B. (2007). *The cost of teacher turnover in five school districts: A pilot study*. Washington, DC: National Commission on Teaching and America's Future. Retrieved March 30, 2009 from http://www.nctaf.org/resources/demonstration_projects/turnover/documents/CTTFullReportfinal.pdf.
- Barnett, W.S. (1996). *Lives in the balance: Age-27 benefit-cost analysis of the High/Scope Perry Preschool Program*. Monographs of the High/Scope Educational Research Foundation, Number Eleven. Ypsilanti, MI: The High/Scope Press.
- Birkeland, S.E. & Curtis, R. (2006). *Ensuring the support and development of new teachers in the Boston Public Schools: A proposal to improve teacher quality and retention*. Boston, MA: Boston Public Schools.
- Boardman, A.E., Greenberg, D.H., Vining, A.R., & Weimer, D.L. (2006). *Cost-benefit analysis: Concepts and practice, 3rd ed.* Upper Saddle River, NJ: Pearson Education.
- Bridges, E.M. (1992). *The incompetent teacher: Managerial responses, a revised and extended edition*. Washington, D.C.: The Falmer Press.
- Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources* 41(Fall): 778-820
- Dawson, T. C., & Billingsley, K. L. (2000). *Unsatisfactory performance: How California's K-12 education system protects mediocrity and how teacher quality can be improved*. San Francisco, CA: Pacific Research Institute for Public Policy.
- Deming, D. (2009). Better schools, less crime? Unpublished manuscript. Retrieved June 8, 2010 from http://sites.google.com/site/demingdavid/Deming_BetterSchoolsLessCrime.pdf?attredirects=0.
- Dillon, S. (2008). Head of teachers' union offers to talk on tenure and merit pay. *New York Times*, November 18. Retrieved March 30, 2009 from <http://www.nytimes.com/2008/11/18/education/18teach.html>.
- Donaldson, M. (2009). So long Lake Wobegon? Using teacher evaluation to improve teacher quality. Washington, DC: Center for American Progress. Retrieved June 8, 2010 from http://www.americanprogress.org/issues/2009/06/pdf/teacher_evaluation.pdf.
- Duncan, A. (2009). Interview with Wolf Blitzer, Transcript. *The Situation Room*. Retrieved March 30, 2009 from <http://transcripts.cnn.com/TRANSCRIPTS/0903/10/sitroom.02.html>.
- Goldstein, J. (2007). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education*, 113, 479-508.
- Goldstein, J. (2004). Making sense of distributed leadership: The case of Peer Assistance and Review. *Educational Evaluation and Policy Analysis* 26 (2): 173-97.

- Honawar, V. (2007). New York City taps lawyers to weed out bad teachers. *Education Week*, 27(14): 13.
- Hummel-Rossi, B., & Ashdown, J. (2002). The state of cost-benefit and cost-effectiveness analyses in education. *Review of Educational Research*, 72(1), 1-30.
- Jackson, C.K. & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4), 85-108.
- Johnson, S.M. & Birkeland, S.E. (2003). Pursuing a 'sense of success': New teachers explain their career decisions. *American Educational Research Journal*, 40(3), 581-617.
- Kemple, J.J. (2008). Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood. New York: MDRC. Retrieved June 8, 2010 from <http://www.mdrc.org/publications/482/full.pdf>.
- Koppich, J. (2005). Addressing teacher quality through induction, professional compensation, and evaluation: The effects on labor-management relations. *Educational Policy*, 19(1): 90-111.
- Koppich, J. (2004). *Toward improving teacher quality: An evaluation of Peer Assistance and Review in Montgomery County Public Schools*. Rockville, MD: Montgomery County Public Schools. Retrieved March 30, 2009 from www.mcps.k12.md.us/departments/development/documents/pgs/PAR_report_final.doc.
- Levin, H.M. (1983). *Cost-effectiveness: A primer*. London: Sage Publications.
- Levin, H.M., & McEwan, P.J. (2001). *Cost-Effectiveness Analysis*. 2nd ed. Thousand Oaks, CA: Sage.
- Lieberman, M. (1998). *Teachers evaluating teachers: Peer review and the new unionism*. New Brunswick, N.J.: Transaction Publishers.
- McEwan, P. J. (2002). Are cost-effectiveness methods used correctly? In *Cost-effectiveness and educational policy*, ed. Henry M. Levin and Patrick J. McEwan, 37-53. Yearbook of the American Education Finance Association. Larchmont, NY: Eye on Education.
- Milanowski, A.T., & Odden, A.R. (2007). A new approach to the cost of teacher turnover. Working Paper 13. Seattle, WA: School Finance Redesign Project, Center on Reinventing Public Education
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook*, 2nd ed. Thousand Oaks, CA: SAGE Publications.
- Murnane, R.J., Willett, J.B. & Levy, F. (1995). "The growing importance of cognitive skills in wage determination." *The Review of Economics and Statistics*, 77(2): 251-66.
- National Commission on Teaching and America's Future. (1996). What matters most: Teaching for America's future. Washington, DC: Author. Retrieved June 8, 2010 from <http://nctaf.org/documents/WhatMattersMost.pdf>.
- Obama, B. (2009, March 26). White House "Open for Questions" Town Hall Transcript. *New York Times*. Retrieved March 30, 2009 from <http://www.nytimes.com/2009/03/26/us/politics/26text-obama.html?pagewanted=all>.

- Rice, J.K. (1997). Cost analysis in education: Paradox and possibility. *Educational Evaluation and Policy Analysis*, 19(4), 309-317.
- Rockoff, J. (2004). The impact of individual teachers on student achievement. Evidence from panel data. *American Economic Review*, 94(2), 247-252.
- Ross, J.A., Barkaoui, K., & Scott, G. (2007). Evaluations that consider the cost of educational programs: The contribution of high-quality studies. *American Journal of Education*, 28(4), 477-492.
- Texas Center for Educational Research. (2000). The cost of teacher turnover. Austin, TX: Author. Retrieved June 8, 2010 from http://www.tcer.org/research/documents/teacher_turnover_full.doc.
- Toch, T. & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Washington, DC: Education Sector.
- Tucker, P. (1997). Lake Wobegon: Where all teachers are competent (Or have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education* 11(2), 103-26.
- U.S. Census Bureau. (2009). *Current Population Survey, Annual Social and Economic Supplements*. Washington, DC: Author. Retrieved June 8, 2010 from <http://www.census.gov/hhes/www/income/data/historical/people/p14AR.xls>.
- van Lier, P. (2008). *Learning from Ohio's best teachers: A homegrown model to improve our schools*. Cleveland, OH: Policy Matters Ohio. Retrieved March 30, 2009 from <http://www.policymattersohio.org/LearningFromOhiosBest.htm>.
- Villar, A. & Strong, M. (2007). Is mentoring worth the money? A benefit-cost analysis and five-year rate of return of a comprehensive mentoring program for beginning teachers. *ERS Spectrum*, 25(3): 1-17.
- Ward, M. E. (1995). Teacher dismissal: The impact of tenure, administrator competence, and other factors. *School Administrator*, 52, 16-19.
- Weingarten, R. (2010). "A New Path Forward: Four Approaches to Quality Teaching and Better Schools," Speech at the National Press Club, January 12. Retrieved June 8, 2010 from <http://www.aft.org/newspubs/news/2010/011210a.cfm>.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The Widget Effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY.: The New Teacher Project. Retrieved June 27, 2009 from <http://widgeteffect.org/>.

Table 1. District Characteristics and Programmatic Features during School Year 2008-09.

	Toledo	Cincinnati	Montgomery County	Rochester	San Juan	Minneapolis	Syracuse
Urbanicity	Urban	Urban	Urban/ Suburb	Urban	Urban/ Suburb	Urban	Urban
Total number of students*	30,423	36,872	139,398	34,096	48,325	38,538	22,123
Total number of teachers*	1,852	2,357	9,371	2,861	2,267	2,250	1,856
Union affiliation	AFT	AFT	NEA	AFT	NEA	NEA/AFT	AFT
Number of years implementing PAR	27	23	6	21	8	24	3
Typical number of CTs per year	10-12	15	30-40	150-200**	6	8	6

*From NCES CCD and district sources.

**CTs work part-time in Rochester.

Table 2. Toledo PAR program costs (2008-09).

Cost Category	Total	Per CT	Per Teacher in PAR	Per Novice*	Per Veteran*
CT Stipends	\$67,650	\$6,150	\$483	\$434	\$867
Substitute teachers for CTs					
- Salary	\$383,471	\$34,861	\$2,739	\$2,458	\$4,916
- Fringe benefits	\$198,341	\$18,031	\$1,417	\$1,271	\$2,543
Substitutes for PAR Panel members**	\$3,960	\$360	\$28	\$25	\$51
CT Training***					
- Training before start of year	\$12,320	\$1,120	\$88	\$79	\$158
- Shadowing of current CT	\$660	\$60	\$5	\$4	\$8
- Attending PAR Panel meetings	\$2,860	\$260	\$20	\$18	\$37
PAR Program Secretary	\$49,249	\$4,477	\$352	\$316	\$631
Administrative costs****					
- Facility	\$30,000	\$2,727	\$214	\$192	\$385
- Computers	\$3,000	\$273	\$21	\$19	\$38
- Mileage reimbursement	\$3,000	\$273	\$21	\$19	\$38
- Supplies	\$1,150	\$105	\$8	\$7	\$15
TOTAL	\$755,661	\$68,592	\$5,398	\$4,844	\$9,688

*In 2008-09, Toledo had 140 total teachers in PAR, 124 novices and 16 veterans. In CT caseload assignment, each veteran counts as two novices; we estimated per-novice and per-veteran resource costs accordingly.

**Three of the five PAR Panel members require substitutes (\$110 per day) for the 12 PAR Panel meeting days.

***Toledo identifies future CTs as "CTs in waiting". In 2008-09, the district had 16 such teachers. They participate in training with the CTs for 2 full days and attend PAR Panel meetings for 2 days during the school year. These costs reflect their daily salary and benefits during training (\$385 per teacher day) and the cost of substitute teachers for the PAR Panel meetings (\$110 per day for 13 CTs in waiting who need substitutes). Six CTs in waiting also shadowed current CTs and required substitutes. Compensation for current CTs is included in their stipend.

****Administrative cost figures represent estimates calculated by the author. Toledo houses its PAR program in a building with several other departments and does not have data on facility cost.

Table 3. PAR Outcomes in the Intervention Program, Average Per Year.

District	Teachers in Intervention	<u>Outcome after Intervention</u>		
		Dismissed	Resigned/Retired	Successful Return to Classroom
Cincinnati	9	5	2	2
Montgomery County	18	5	4	9
Rochester	2	0	0	2
San Juan	3	1	1	2
Toledo	3	--	1	1

NOTE: Data self-reported by the district PAR program.

Table 4. The costs and benefits of teacher Peer Assistance and Review.

	Costs of PAR	Benefits of PAR
Financial	<ul style="list-style-type: none"> • \$3,000 to \$7,000 per novice teacher served • \$6,000 to \$10,000 per veteran teacher served <ul style="list-style-type: none"> - Salaries, benefits, and training costs for replacement teachers for CTs - CT stipends (\$5,000 to \$10,000 per CT) - PAR Panel stipends - Substitute coverage for PAR Panel - CT and PAR Panel training - Program management costs (e.g., program director) - Other administrative costs 	<ul style="list-style-type: none"> • Cost-savings from reduced teacher turnover • Cost-savings from reduced dismissals • Long-term cost-savings from improved teacher quality
Organizational	<ul style="list-style-type: none"> • Removes some of the districts' best teachers from the classroom for several years. 	<ul style="list-style-type: none"> • Increases teacher performance in the classroom through support and evaluation • Improves organizational stability through improved retention • Builds a positive organizational culture centered on instruction • Alleviates burden on principals of evaluating novice and underperforming veteran teachers • Creates career opportunities and professional development for CTs • Improves labor-management relations at the district and school levels